

1/14



OIPE

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:55

Input Set : N:\Crf3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw

1 >110> APPLICANT: De Robertis, Edward M.
 2 Bouwmeester, Tewis
 3 >120> TITLE OF INVENTION: Endoderm, Cardiac and Neural Inducing
 4 Factors
 5 >130> FILE REFERENCE: 510015-248
 6 >140> CURRENT APPLICATION NUMBER: US/09/903,180A
 7 >141> CURRENT FILING DATE: 2001-07-11
 8 >150> PRIOR APPLICATION NUMBER: 09/552,988
 9 >151> PRIOR FILING DATE: 2000-04-21
 10 >150> PRIOR APPLICATION NUMBER: US 08/878,474
 11 >151> PRIOR FILING DATE: 1997-06-18
 12 >150> PRIOR APPLICATION NUMBER: US 60/020,150
 13 >151> PRIOR FILING DATE: 1996-06-20
 14 >160> NUMBER OF SEQ ID NOS: 10
 15 >170> SOFTWARE: FastSEQ for Windows Version 3.0
 17 >210> SEQ ID NO: 1
 18 >211> LENGTH: 270
 19 >212> TYPE: PRT
 20 >213> ORGANISM: Xenopus
 21 >400> SEQUENCE: 1
 22 Met Leu Leu Asn Val Leu Arg Ile Cys Ile Ile Val Cys Leu Val Asn
 23 1 5 10 15
 24 Asp Gly Ala Gly Lys His Ser Glu Gly Arg Glu Arg Thr Lys Thr Tyr
 25 20 25 30
 26 Ser Leu Asn Ser Arg Gly Tyr Phe Arg Lys Glu Arg Gly Ala Arg Arg
 27 35 40 45
 28 Ser Lys Ile Leu Leu Val Asn Thr Lys Gly Leu Asp Glu Pro His Ile
 29 50 55 60
 30 Gly His Gly Asp Phe Gly Leu Val Ala Glu Leu Phe Asp Ser Thr Arg
 31 65 70 75 80
 32 Thr His Thr Asn Arg Lys Glu Pro Asp Met Asn Lys Val Lys Leu Phe
 33 85 90 95
 34 Ser Thr Val Ala His Gly Asn Lys Ser Ala Arg Arg Lys Ala Tyr Asn
 35 100 105 110
 36 Gly Ser Arg Arg Asn Ile Phe Ser Arg Arg Ser Phe Asp Lys Arg Asn
 37 115 120 125
 38 Thr Glu Val Thr Glu Lys Pro Gly Ala Lys Met Phe Trp Asn Asn Phe
 39 130 135 140
 40 Leu Val Lys Met Asn Gly Ala Pro Gln Asn Thr Ser His Gly Ser Lys
 41 145 150 155 160
 42 Ala Gln Glu Ile Met Lys Glu Ala Cys Lys Thr Leu Pro Phe Thr Gln
 43 165 170 175
 44 Asn Ile Val His Glu Asn Cys Asp Arg Met Val Ile Gln Asn Asn Leu

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:55

Input Set : N:\Crf3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw

```

45          180          185          190
46 Cys Phe Gly Lys Cys Ile Ser Leu His Val Pro Asn Gln Gln Asp Arg
47          195          200          205
48 Arg Asn Thr Cys Ser His Cys Leu Pro Ser Lys Phe Thr Leu Asn His
49          210          215          220
50 Leu Thr Leu Asn Cys Thr Gly Ser Lys Asn Val Val Lys Val Val Met
51          225          230          235          240
52 Met Val Glu Glu Cys Thr Cys Glu Ala His Lys Ser Asn Phe His Gln
53          245          250          255
54 Thr Ala Gln Phe Asn Met Asp Thr Ser Thr Thr Leu His His
55          260          265          270
57 <210> SEQ ID NO: 2
58 <211> LENGTH: 1411
59 <212> TYPE: DNA
60 <213> ORGANISM: Xenopus
61 <400> SEQUENCE: 2
62 gaattcctaa aagcggcaca gtgcaggaac agcaagtgcg tcagaaacac tgcaggggtct 60
63 agatatcata caatgttaact aaatgtactc aggatctgta ttatgtctcg ctttgtgaat 120
64 gatggagcag gaaaacactc agaaggacga gaaaggacaa aaacatattc acttaacago 180
65 agaggttact tcagaaaaga aagaggagca cgtaggagca agattctgct ggtgaatact 240
66 aaaggtcttg atgaacccca cattgggcat ggtgattttg gcttagtagc tgaactatct 300
67 gattccacca gaacacatac aaacagaaaa gagccagaca tgaacaaagt caagcttttc 360
68 tcaacagttg cccatggaaa caaaagtgcg agaagaaaag cttacaatgg ttctagaagg 420
69 aatatttttt ctgcgcgttc ttttgataaa agaaatacag aggttaactga aaagcctggg 480
70 gccaaagatg tctggaacaa ttttttggtt aaaatgaatg gagcccaaca gaatacaagc 540
71 catggcagta aagcacagga aataatgaaa gaagcttgca aaaccttgcc ctccactcag 600
72 aatattgtac atgaaaaactg tgacaggatg gtgatacaga acaatctgtg ctttggtaaa 660
73 tgcactctct tccatgttcc aaatcagcaa gatcgacgaa ataactgttc ccattgcttg 720
74 ccgtccaaat ttacctgaa ccacctgacg ctgaattgta ctggatctaa gaatgtagta 780
75 aaggttgta tgaatgtaga ggaatgcacg tgtgaagctc ataagagcaa ctccacccaa 840
76 actgcacagt ttaacatgga tacatctact acctgcacc attaaaagga ctgtctgcca 900
77 tacagtatgg aaatgcccat ttgttggaat attogttaca tgcctatgat ctaaagcatt 960
78 atgttgccct ctgtttcata taaccacatg gaataaggat tgtatgaatt ataattaaca 1020
79 aatggcattt tgtgtaacat gcaagatctc tgttccatca gttgcaagat aaaaggcaat 1080
80 atttgtttga cttttttcta caaaatgaat acccaaatat atgataagat aatggggta 1140
81 aaactgttaa ggggtaatgt aataataggg actaacaacc aatcagcagg tatgatttac 1200
82 tggtaacctg tttaaaagca aacatcttat tggttgctat gggttactgc ttctgggcaa 1260
83 aatgtgtgcc tcataggggg gttagtgtgt tgtgtactga attaattgta tttatttcac 1320
84 tgttacaatg aagaggatgt ctatgtttat ttcactttta ttaatgtaca ataaatgttc 1380
85 ttgtttcttt aaaaaaaaaa aaaaactcga g 1411
87 <210> SEQ ID NO: 3
88 <211> LENGTH: 318
89 <212> TYPE: PRT
90 <213> ORGANISM: Xenopus
91 <400> SEQUENCE: 3
92 Met Ser Arg Thr Arg Lys Val Asp Ser Leu Leu Leu Leu Ala Ile Pro
93 1 5 10 15
94 Gly Leu Ala Leu Leu Leu Leu Pro Asn Ala Tyr Cys Ala Ser Cys Glu
95 20 25 30

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:55

Input Set : N:\Crif3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw

```

96   Pro Val Arg Ile Pro Met Cys Lys Ser Met Pro Trp Asn Met Thr Lys
97           35                               40                   45
98   Met Pro Asn His Leu His His Ser Thr Gln Ala Asn Ala Ile Leu Ala
99           50                               55                   60
100  Ile Glu Gln Phe Glu Gly Leu Leu Thr Thr Glu Cys Ser Gln Asp Leu
101           65                               70                   75                   80
102  Leu Phe Phe Leu Cys Ala Met Tyr Ala Pro Ile Cys Thr Ile Asp Phe
103           85                               90                   95
104  Gln His Glu Pro Ile Lys Pro Cys Lys Ser Val Cys Glu Arg Ala Arg
105           100                              105                   110
106  Ala Gly Cys Glu Pro Ile Leu Ile Lys Tyr Arg His Thr Trp Pro Glu
107           115                              120                   125
108  Ser Leu Ala Cys Glu Glu Leu Pro Val Tyr Asp Arg Gly Val Cys Ile
109           130                              135                   140
110  Ser Pro Glu Ala Ile Val Thr Val Glu Gln Gly Thr Asp Ser Met Pro
111           145                              150                   155                   160
112  Asp Phe Ser Met Asp Ser Asn Asn Gly Asn Cys Gly Ser Gly Arg Glu
113           165                              170                   175
114  His Cys Lys Cys Lys Pro Met Lys Ala Thr Gln Lys Thr Tyr Leu Lys
115           180                              185                   190
116  Asn Asn Tyr Asn Tyr Val Ile Arg Ala Lys Val Lys Glu Val Lys Val
117           195                              200                   205
118  Lys Cys His Asp Ala Thr Ala Ile Val Glu Val Lys Glu Ile Leu Lys
119           210                              215                   220
120  Ser Ser Leu Val Asn Ile Pro Lys Asp Thr Val Thr Leu Tyr Thr Asn
121           225                              230                   235                   240
122  Ser Gly Cys Leu Cys Pro Gln Leu Val Ala Asn Glu Glu Tyr Ile Ile
123           245                              250                   255
124  Met Gly Tyr Glu Asp Lys Glu Arg Thr Arg Leu Leu Leu Val Glu Gly
125           260                              265                   270
126  Ser Leu Ala Glu Lys Trp Arg Asp Arg Leu Ala Lys Lys Val Lys Arg
127           275                              280                   285
128  Trp Asp Gln Lys Leu Arg Arg Pro Arg Lys Ser Lys Asp Pro Val Ala
129           290                              295                   300
130  Pro Ile Pro Asn Lys Asn Ser Asn Ser Arg Gln Ala Arg Ser
131           305                              310                   315

```

133 (210) SEQ ID NO: 4

134 (211) LENGTH: 1875

135 (212) TYPE: DNA

136 (213) ORGANISM: Xenopus

137 (400) SEQUENCE: 4

```

138   gaattccctt tcacacagga ctcttgccag aggtgaatgg ttagccctat ggatttggtt      60
139   tgttgatttt gacacatgat tgattgcttt cagataggat tgaaggactt ggatttttat      120
140   ctaattctgc acttttaa attatctgagta attgttcatt ttgtattgga tgggactaaa      180
141   gataaactta actccttgct ttgacttgcc ccataaacta taagggtgggg tgagttgtag      240
142   ttgctttttac atgtgccagc attttccctg tattccctgt attccctcta aagtaagcct      300
143   acacatacag gttgggcaga ataacaatgt ctogaacaag gaaagtggac tcattactgc      360
144   tactggccat acctggactg gcgcttctct tattaaccaa tgcttactgt gcttcgtgtg      420
145   agcctgtgcg gatcccatg tgcaaatcta tgccatggaa catgaccaag atgcccaccc      480

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:55

Input Set : N:\Crf3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw

```

146 atctccacca cagcactcaa gccaatgcca tcttgccaat tgaacagttt gaaggtttgc 540
147 tgaccactga atgtagccag gaccttttgt tctttctgtg tgccatgtat gccccattt 600
148 gtaccatcga ttccagcat gaaccaatta agccttgcaa gtccgtgtgc gaaagggcca 660
149 gggccggctg tgagcccat ctcataaagt accggcacac ttggccagag agcctggcat 720
150 gtgaagagct gcccgtatat gacagaggag tctgcattct cccagaggct atcgtcacag 780
151 tggaacaagg aacagattca atgccagact tctccatgga ttcaacaat ggaaattgcg 840
152 gaagcggcag ggagcactgt aaatgcaagc ccattgaagg aacccaaaag acgtatctca 900
153 agaataatta caattatgta atcagagcaa aagtgaagga ggtgaaagt aaatgcacg 960
154 acgcaacagc aattgtggaa gtaaaggaga ttctcaagtc ttccctagtg aacattctca 1020
155 aagcacacagt gacactgtac accaactcag gctgcttggt cccccagctt gttgccaatg 1080
156 aggaatacat aattatgggc tatgaagaca aagagcgtac caggtctcta ctagtggag 1140
157 gatccttggc cgaaaaatgg agagatcgtc ttgctaagaa agtcaagcgc tgggatcaaa 1200
158 agcttcgacg tcccaggaag agcaaaagacc ccgtggctcc aattcccaac aaaaacagca 1260
159 attccagaca agcgcgtagt tagactaacg gaaagggtga tggaaactct atggactttg 1320
160 aaactaagat ttgcattgtt ggaagagcaa aaaagaaatt gcactacagc acgttatatt 1380
161 ctattgttta ctacaagaag ctggtttagt tgattgtagt tctcctttcc ttcttttttt 1440
162 ttataactat atttgcacgt gttcccaggc aattgtttta ttcaacttcc agtgacagag 1500
163 cagtgaactga atgtctcagc ctaaaagaag tcaattcatt tctgatcaac taatggtgac 1560
164 aagtgtttga tacttgggga aagtgaacta attgcaatgg taaatcagag aaaagttgac 1620
165 caatgtttgt ttctctgtag atgaacaagt gagagatcac atttaaatga tgatcacttt 1680
166 ccatttaata ctttcagcag ttttagttag atgacatgta ggatgcacct aaatctaaat 1740
167 attttatcat aaatgaagag ctggtttaga ctgtatggtc actgttggga aggtaaatgc 1800
168 ctactttgtc aattctgttt taaaaattgc ctaaaataat attaatgctt aaataaaaaa 1860
169 aaaaaaaaaa aaaaa 1875

```

171 <210> SEQ ID NO: 5

172 <211> LENGTH: 979

173 <212> TYPE: PRT

174 <213> ORGANISM: Xenopus

175 <400> SEQUENCE: 5

```

176 Met Leu Leu Leu Phe Arg Ala Ile Pro Met Leu Leu Leu Gly Leu Met
177 1 5 10 15
178 Val Leu Gln Thr Asp Cys Glu Ile Ala Gln Tyr Tyr Ile Asp Glu Glu
179 20 25 30
180 Glu Pro Pro Gly Thr Val Ile Ala Val Leu Ser Gln His Ser Ile Phe
181 35 40 45
182 Asn Thr Thr Asp Ile Pro Ala Thr Asn Phe Arg Leu Met Lys Gln Phe
183 50 55 60
184 Asn Asn Ser Leu Ile Gly Val Arg Glu Ser Asp Gly Gln Leu Ser Ile
185 65 70 75 80
186 Met Glu Arg Ile Asp Arg Glu Gln Ile Cys Arg Gln Ser Leu His Cys
187 85 90 95
188 Asn Leu Ala Leu Asp Val Val Ser Phe Ser Lys Gly His Phe Lys Leu
189 100 105 110
190 Leu Asn Val Lys Val Glu Val Arg Asp Ile Asn Asp His Ser Pro His
191 115 120 125
192 Phe Pro Ser Glu Ile Met His Val Glu Val Ser Glu Ser Ser Ser Val
193 130 135 140
194 Gly Thr Arg Ile Pro Leu Glu Ile Ala Ile Asp Glu Asp Val Gly Ser
195 145 150 155 160

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:55

Input Set : N:\Crf3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw

196	Asn	Ser	Ile	Gln	Asn	Phe	Gln	Ile	Ser	Asn	Asn	Ser	His	Phe	Ser	Ile
197					165					170					175	
198	Asp	Val	Leu	Thr	Arg	Ala	Asp	Gly	Val	Lys	Tyr	Ala	Asp	Leu	Val	Leu
199				180					185					190		
200	Met	Arg	Glu	Leu	Asp	Arg	Glu	Ile	Gln	Pro	Thr	Tyr	Ile	Met	Glu	Leu
201			195					200					205			
202	Leu	Ala	Met	Asp	Gly	Gly	Val	Pro	Ser	Leu	Ser	Gly	Thr	Ala	Val	Val
203		210					215					220				
204	Asn	Ile	Arg	Val	Leu	Asp	Phe	Asn	Asp	Asn	Ser	Pro	Val	Phe	Glu	Arg
205	225				230						235					240
206	Ser	Thr	Ile	Ala	Val	Asp	Leu	Val	Glu	Asp	Ala	Pro	Leu	Gly	Tyr	Leu
207				245					250						255	
208	Leu	Leu	Glu	Leu	His	Ala	Thr	Asp	Asp	Asp	Glu	Gly	Val	Asn	Gly	Glu
209				260					265					270		
210	Ile	Val	Tyr	Gly	Phe	Ser	Thr	Leu	Ala	Ser	Gln	Glu	Val	Arg	Gln	Leu
211			275					280					285			
212	Phe	Lys	Ile	Asn	Ser	Arg	Thr	Gly	Ser	Val	Thr	Leu	Glu	Gly	Gln	Val
213		290					295					300				
214	Asp	Phe	Glu	Thr	Lys	Gln	Thr	Tyr	Glu	Phe	Glu	Val	Gln	Ala	Gln	Asp
215	305				310						315					320
216	Leu	Gly	Pro	Asn	Pro	Leu	Thr	Ala	Thr	Cys	Lys	Val	Thr	Val	His	Ile
217				325						330					335	
218	Leu	Asp	Val	Asn	Asp	Asn	Thr	Pro	Ala	Ile	Thr	Ile	Thr	Pro	Leu	Thr
219			340					345						350		
220	Thr	Val	Asn	Ala	Gly	Val	Ala	Tyr	Ile	Pro	Glu	Thr	Ala	Thr	Lys	Glu
221			355					360					365			
222	Asn	Phe	Ile	Ala	Leu	Ile	Ser	Thr	Thr	Asp	Arg	Ala	Ser	Gly	Ser	Asn
223		370					375					380				
224	Gly	Gln	Val	Arg	Cys	Thr	Leu	Tyr	Gly	His	Glu	His	Phe	Lys	Leu	Gln
225	385				390						395					400
226	Gln	Ala	Tyr	Glu	Asp	Ser	Tyr	Met	Ile	Val	Thr	Thr	Ser	Thr	Leu	Asp
227				405						410					415	
228	Arg	Glu	Asn	Ile	Ala	Ala	Tyr	Ser	Leu	Thr	Val	Val	Ala	Glu	Asp	Leu
229			420						425					430		
230	Gly	Phe	Pro	Ser	Leu	Lys	Thr	Lys	Lys	Tyr	Tyr	Thr	Val	Lys	Val	Ser
231		435						440					445			
232	Asp	Glu	Asn	Asp	Asn	Ala	Pro	Val	Phe	Ser	Lys	Pro	Gln	Tyr	Glu	Ala
233		450					455					460				
234	Ser	Ile	Leu	Glu	Asn	Asn	Ala	Pro	Gly	Ser	Tyr	Ile	Thr	Thr	Val	Ile
235	465				470						475					480
236	Ala	Arg	Asp	Ser	Asp	Ser	Asp	Gln	Asn	Gly	Lys	Val	Asn	Tyr	Arg	Leu
237				485						490					495	
238	Val	Asp	Ala	Lys	Val	Met	Gly	Gln	Ser	Leu	Thr	Thr	Phe	Val	Ser	Leu
239			500						505					510		
240	Asp	Ala	Asp	Ser	Gly	Val	Leu	Arg	Ala	Val	Arg	Ser	Leu	Asp	Tyr	Glu
241		515						520					525			
242	Lys	Leu	Lys	Gln	Leu	Asp	Phe	Glu	Ile	Glu	Ala	Ala	Asp	Asn	Gly	Ile
243		530					535					540				
244	Pro	Gln	Leu	Ser	Thr	Arg	Val	Gln	Leu	Asn	Leu	Arg	Ile	Val	Asp	Gln

VERIFICATION SUMMARY

DATE: 02/06/2002

PATENT APPLICATION: US/09/903,180A

TIME: 09:21:56

Input Set : N:\Crf3\RULE60\09903180A.raw

Output Set: N:\CRF3\02062002\I903180A.raw